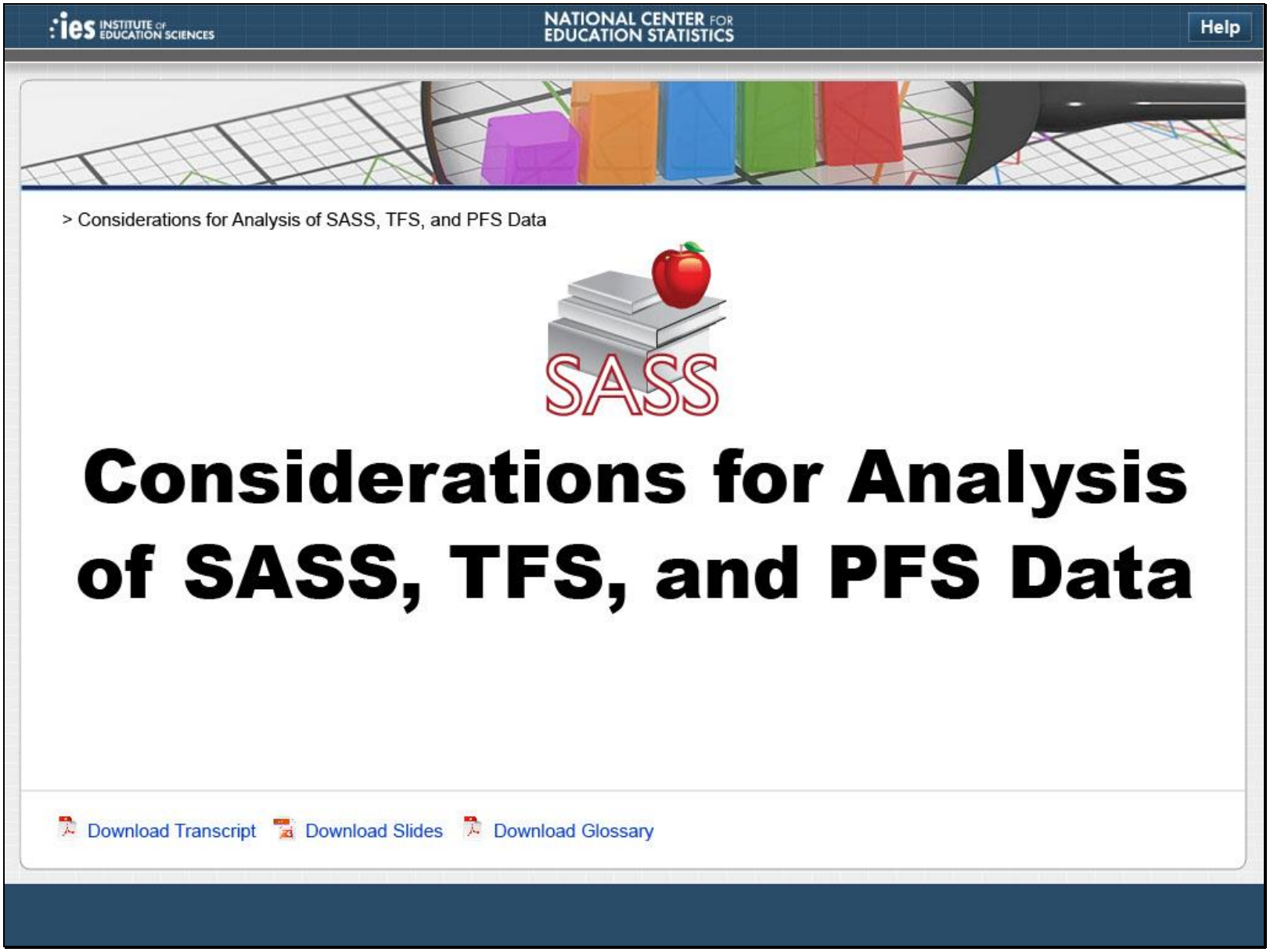
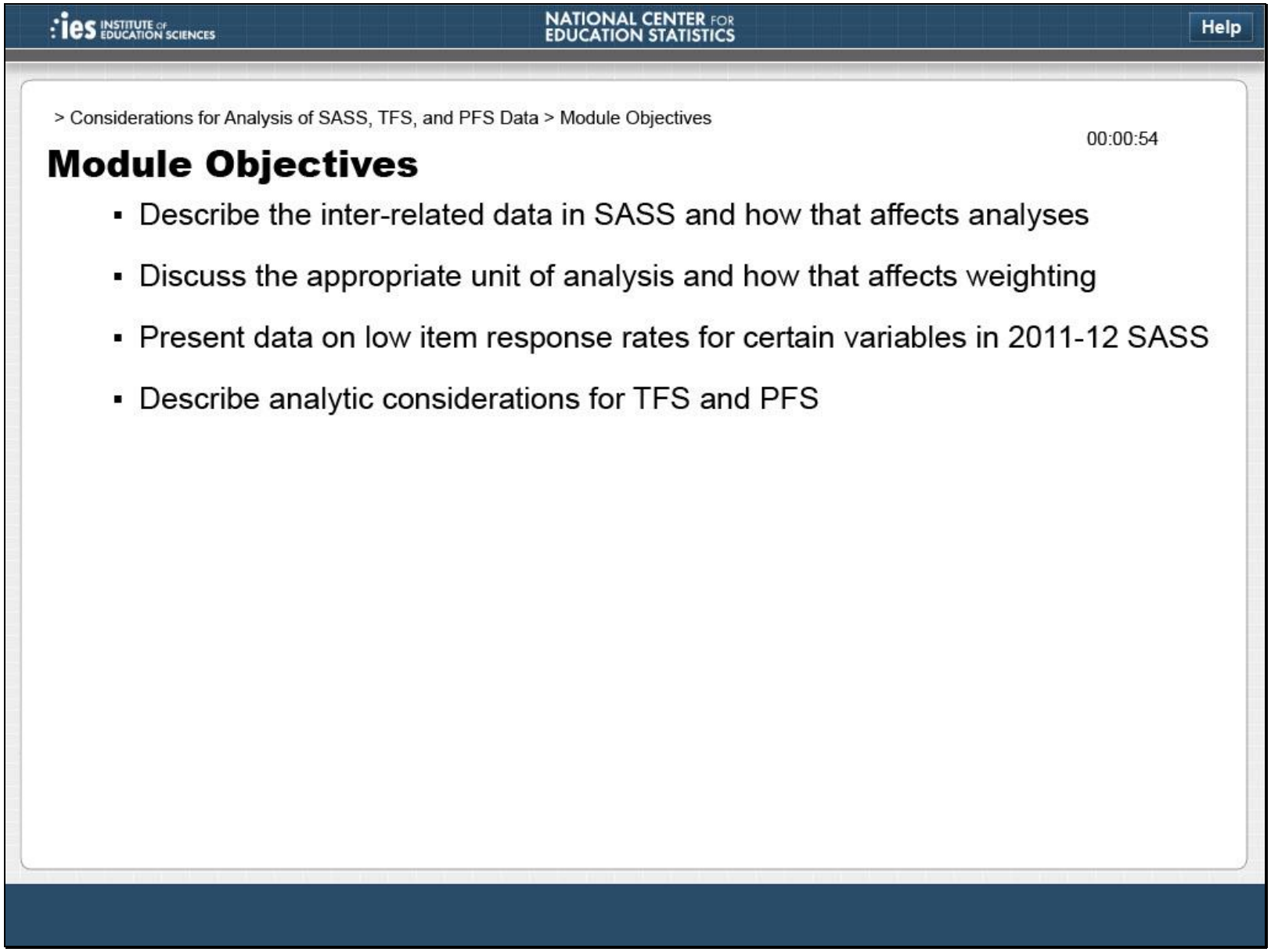
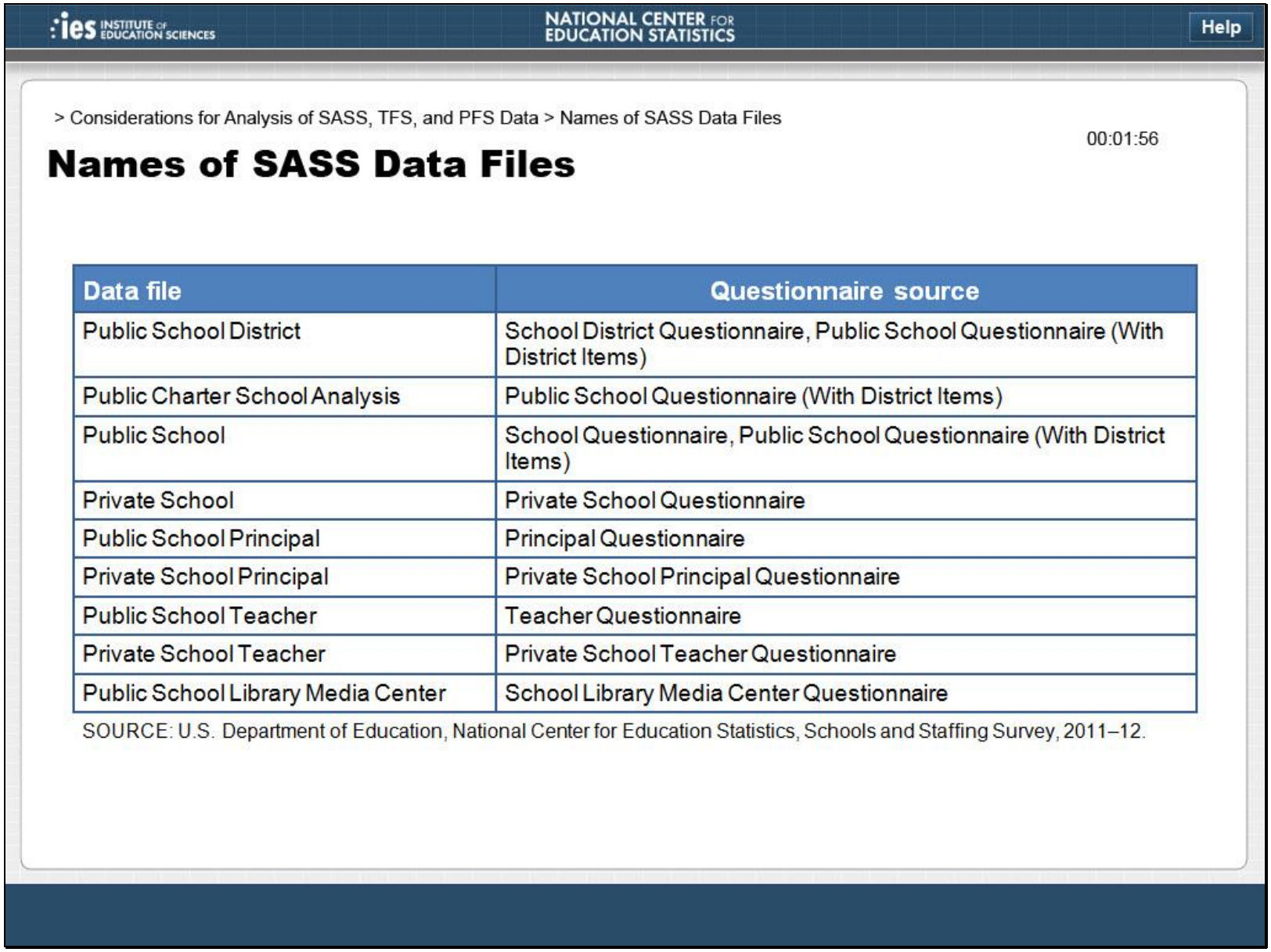
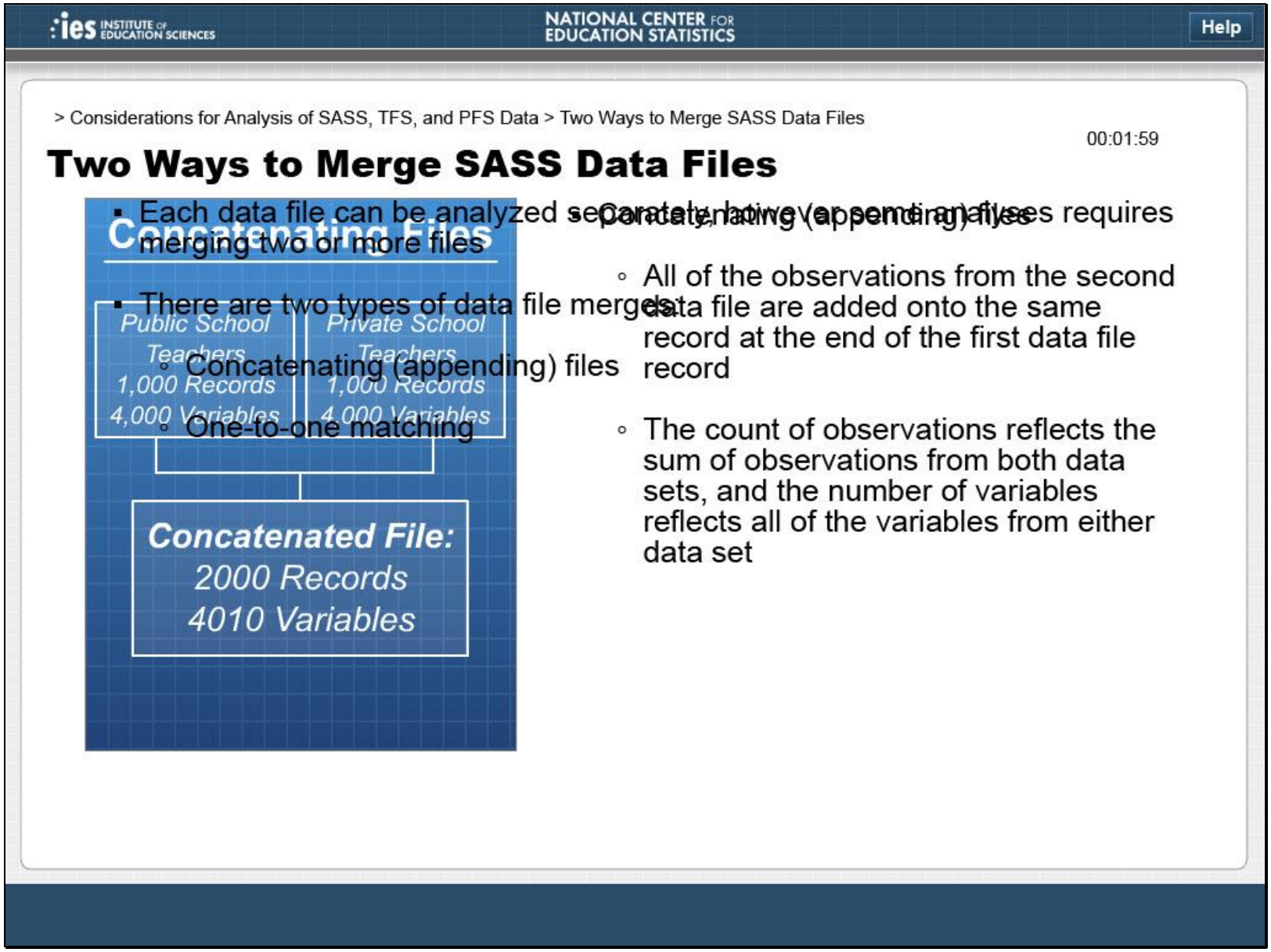

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data
SASS
Considerations for Analysis of SASS, TFS, and PFS Data
Download Transcript
Download Slides
Download Glossary


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Module Objectives
00:00:54
Module Objectives
Describe the inter-related data in SASS and how that affects analyses
Discuss the appropriate unit of analysis and how that affects weighting
Present data on low item response rates for certain variables in 2011-12 SASS
Describe analytic considerations for TFS and PFS



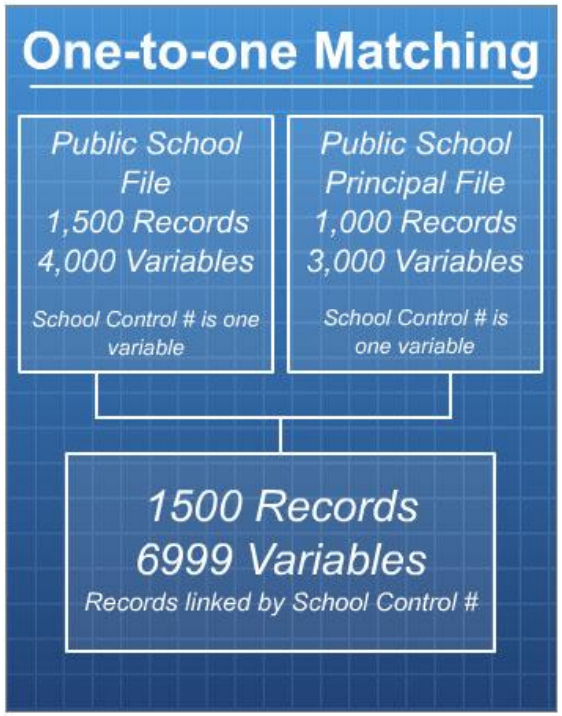
Data file	Questionnaire source
Public School District	School District Questionnaire, Public School Questionnaire (With District Items)
Public Charter School Analysis	Public School Questionnaire (With District Items)
Public School	School Questionnaire, Public School Questionnaire (With District Items)
Private School	Private School Questionnaire
Public School Principal	Principal Questionnaire
Private School Principal	Private School Principal Questionnaire
Public School Teacher	Teacher Questionnaire
Private School Teacher	Private School Teacher Questionnaire
Public School Library Media Center	School Library Media Center Questionnaire

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey, 2011–12.

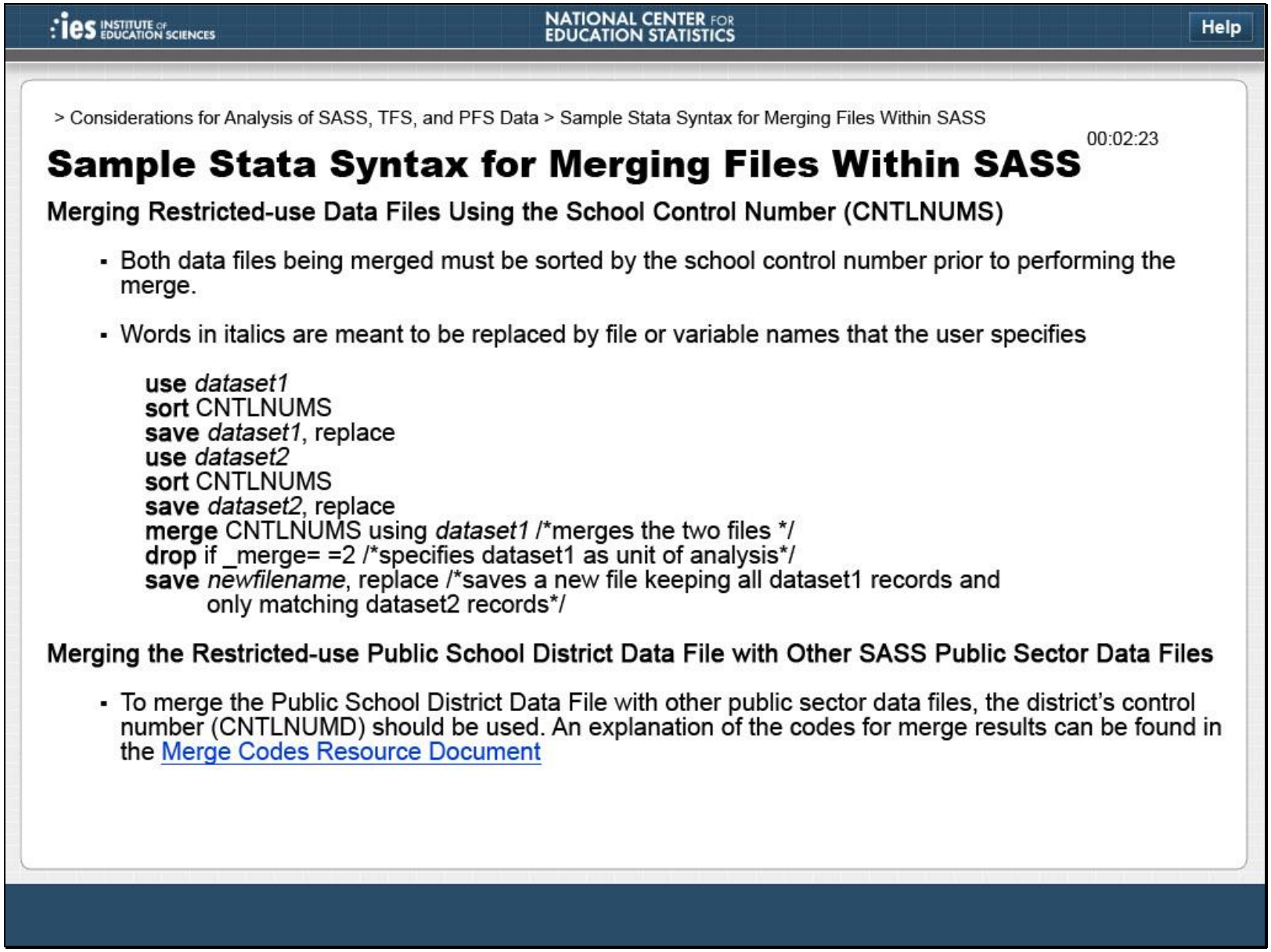

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Two Ways to Merge SASS Data Files
00:01:59
Two Ways to Merge SASS Data Files
Each data file can be analyzed separately; however, some analyses requires merging two or more files
There are two types of data file merges:
Concatenating (appending) files
One-to-one matching
Concatenating Files
Public School Teachers
1,000 Records
4,000 Variables
Private School Teachers
1,000 Records
4,000 Variables
Concatenated File:
2000 Records
4010 Variables
Concatenating (appending) files
All of the observations from the second data file are added onto the same record at the end of the first data file record
The count of observations reflects the sum of observations from both data sets, and the number of variables reflects all of the variables from either data set

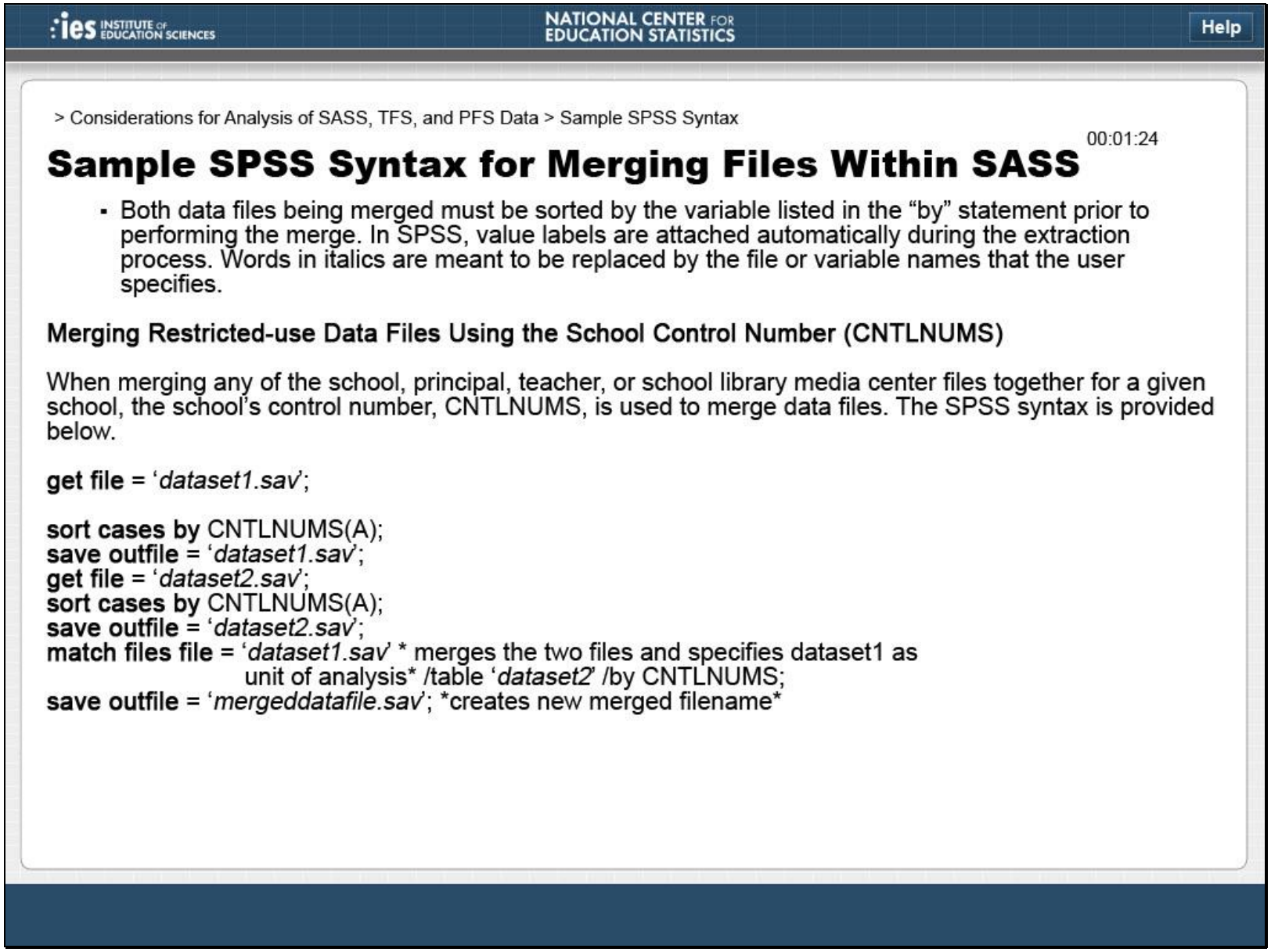
00:01:31

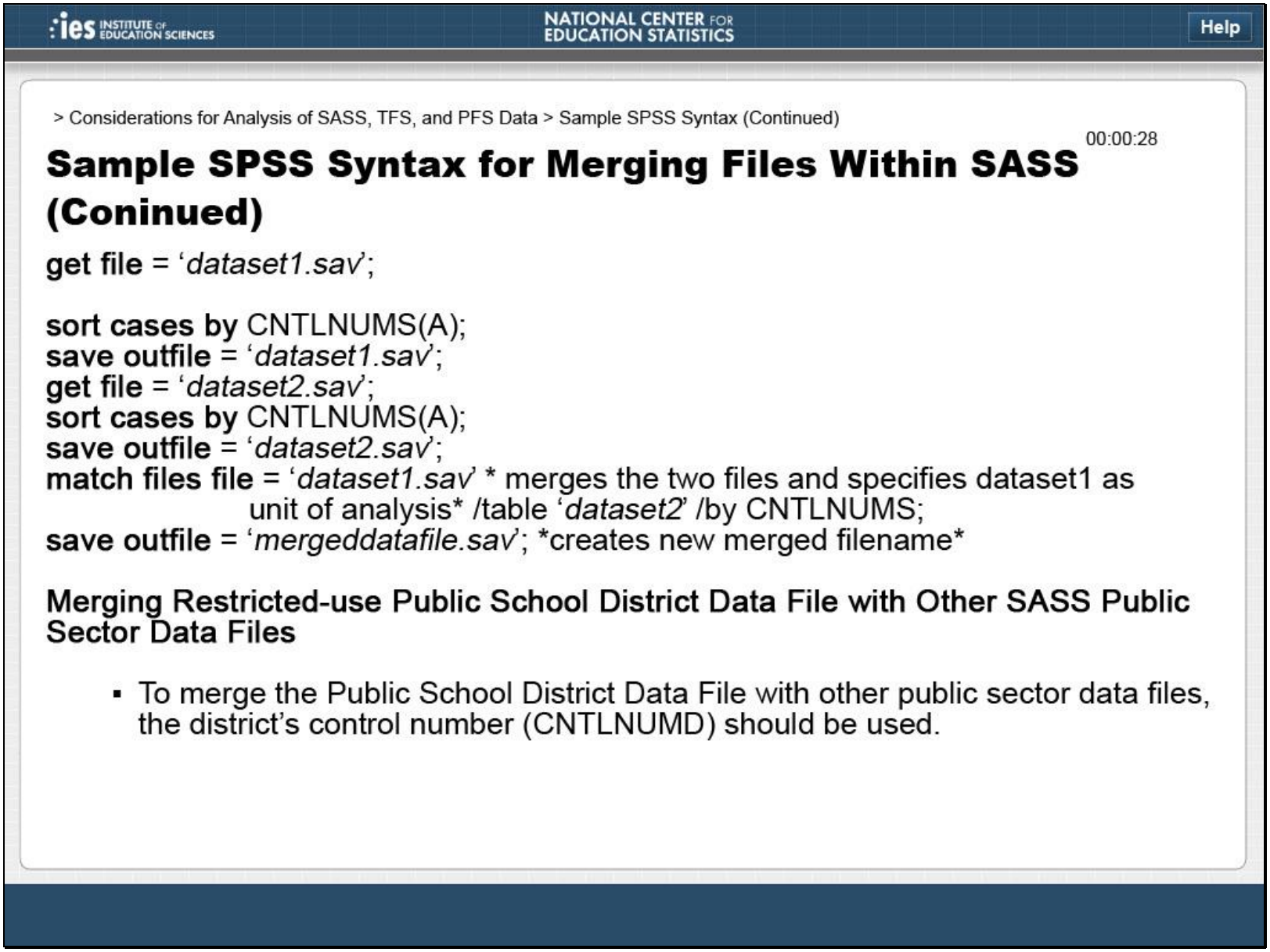
Two Ways to Merge SASS Data Files (Continued)

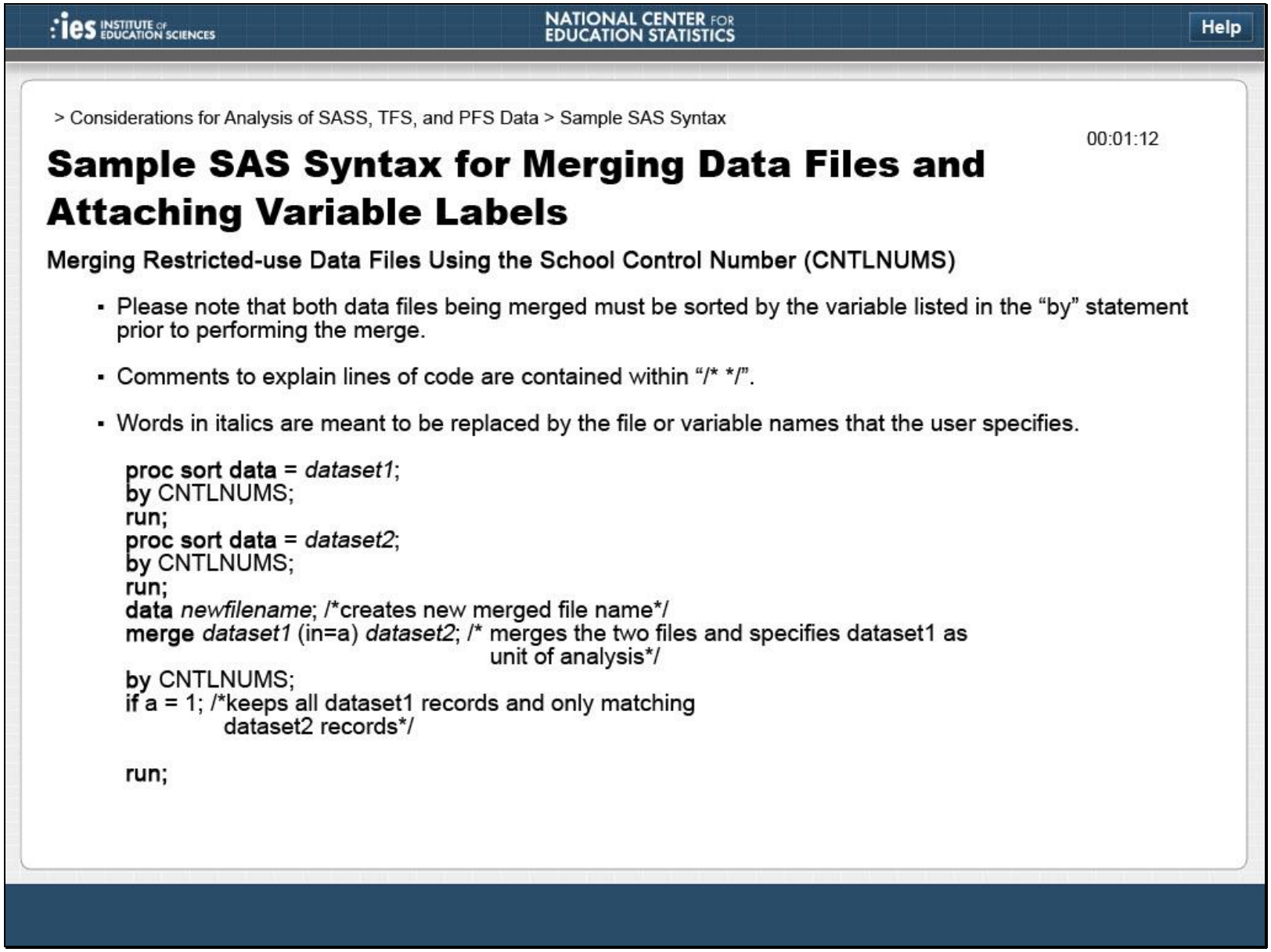


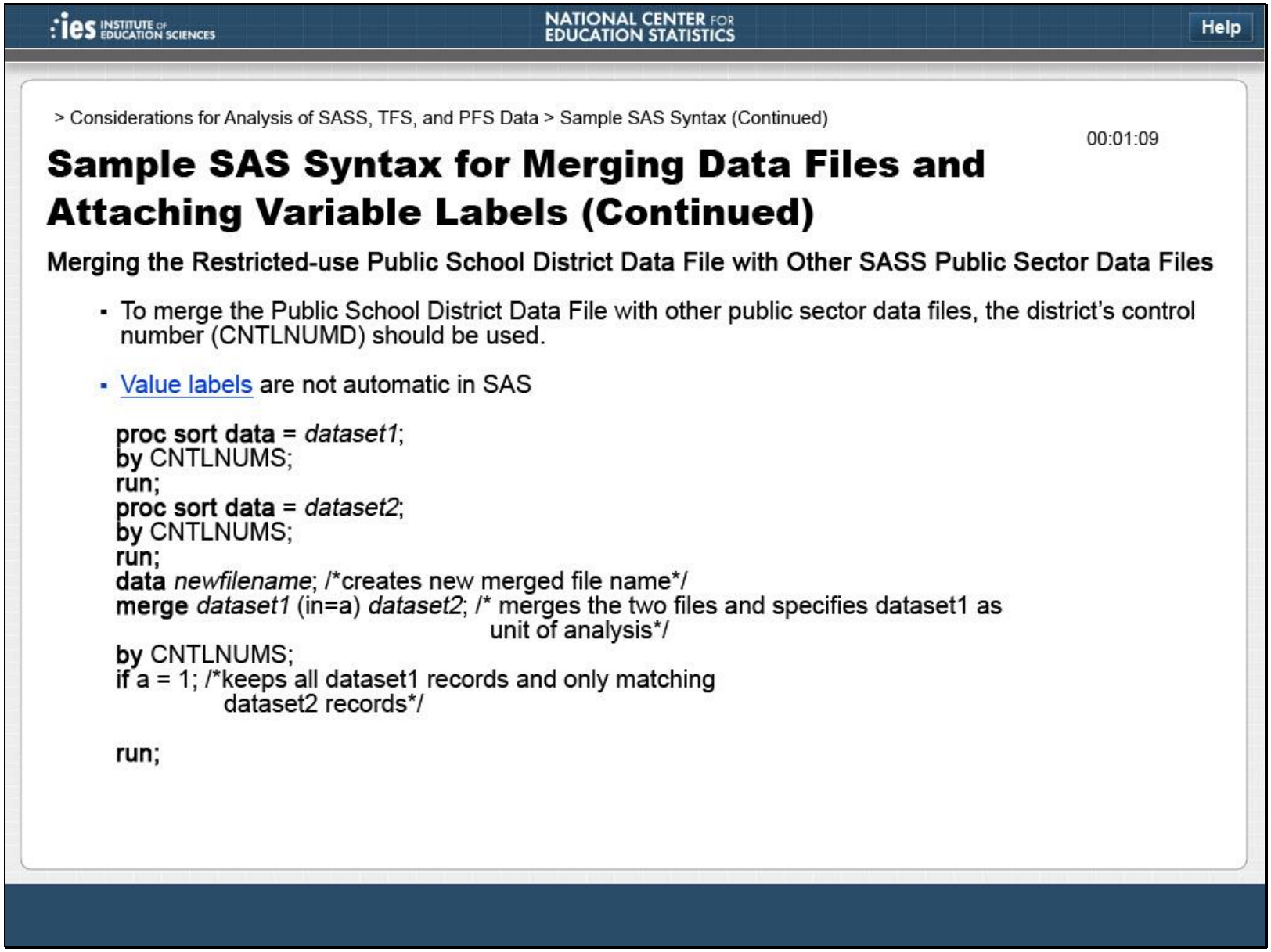
- One-to-one matching
  - Each data record is matched by a unique identifying variable
  - The variables from the second data file that do not match are added onto the matched record from the first data file
  - Matched data records will have additional data in variables that match

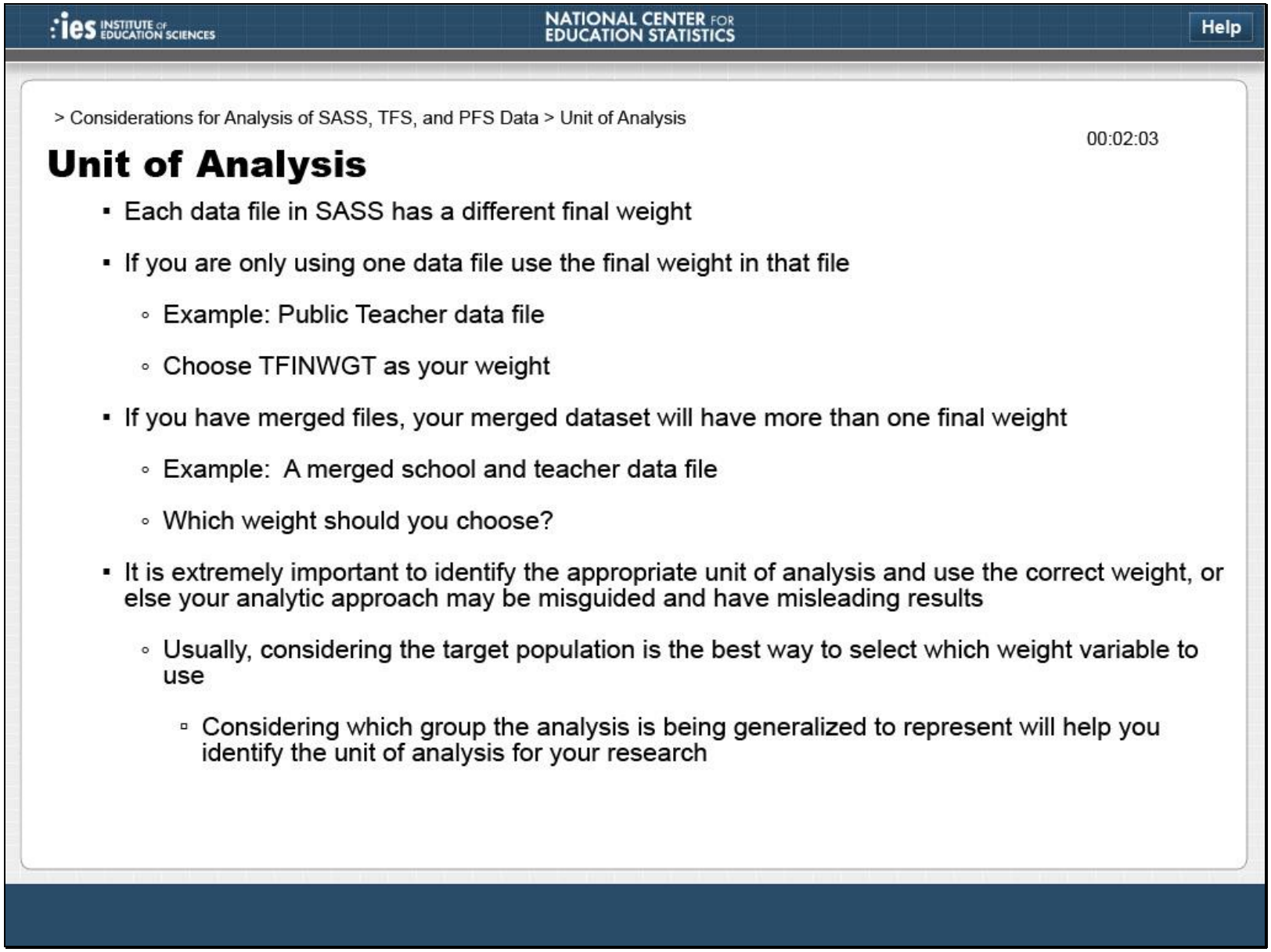

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Sample Stata Syntax for Merging Files Within SASS
00:02:23
Sample Stata Syntax for Merging Files Within SASS
Merging Restricted-use Data Files Using the School Control Number (CNTLNUMS)
Both data files being merged must be sorted by the school control number prior to performing the merge.
Words in italics are meant to be replaced by file or variable names that the user specifies
use dataset1
sort CNTLNUMS
save dataset1, replace
use dataset2
sort CNTLNUMS
save dataset2, replace
merge CNTLNUMS using dataset1 /*merges the two files */
drop if _merge= =2 /*specifies dataset1 as unit of analysis*/
save newfilename, replace /*saves a new file keeping all dataset1 records and only matching dataset2 records*/
Merging the Restricted-use Public School District Data File with Other SASS Public Sector Data Files
To merge the Public School District Data File with other public sector data files, the district's control number (CNTLNUMD) should be used. An explanation of the codes for merge results can be found in the Merge Codes Resource Document

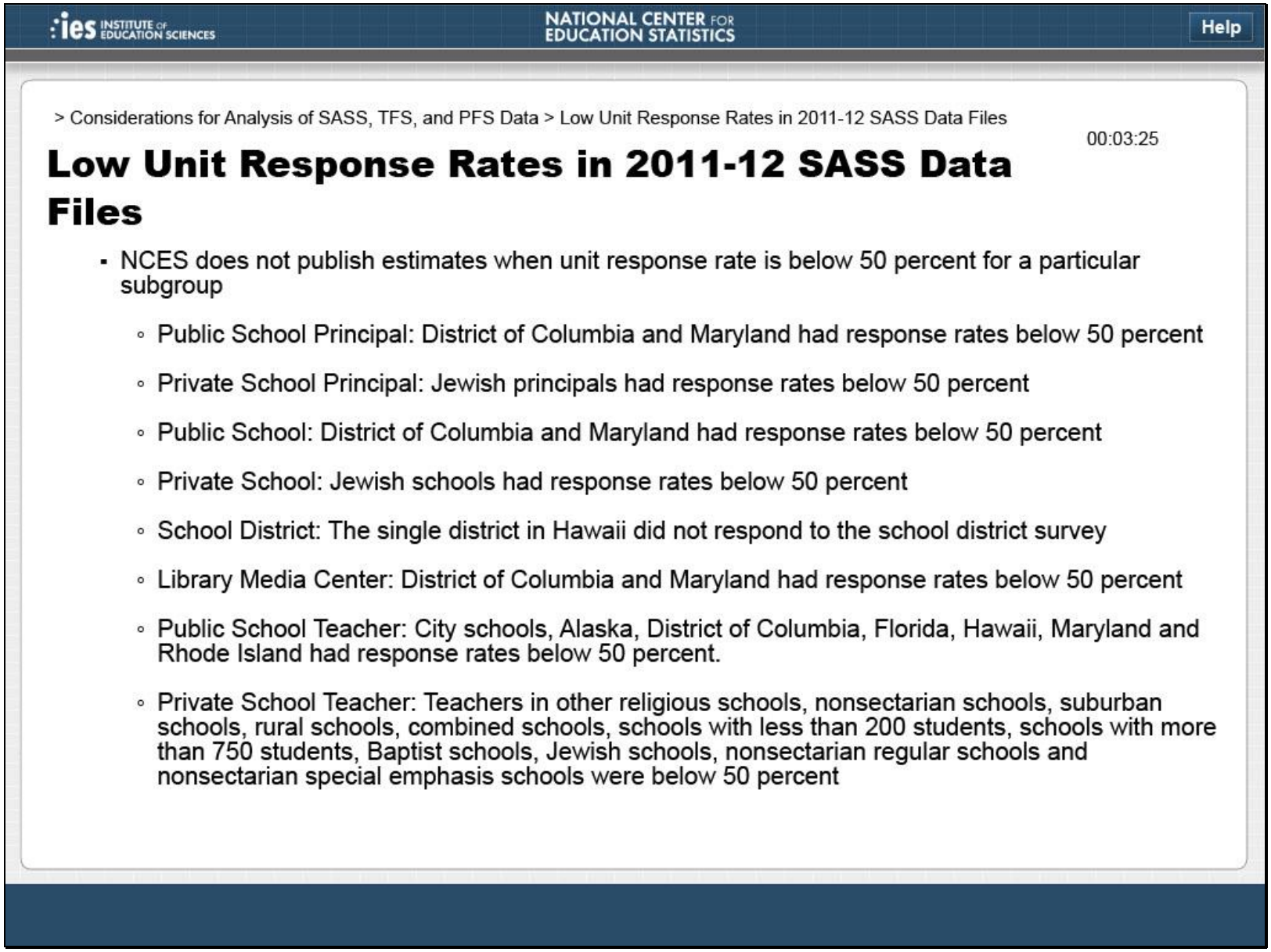

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Sample SPSS Syntax
00:01:24
Sample SPSS Syntax for Merging Files Within SASS
Both data files being merged must be sorted by the variable listed in the "by" statement prior to performing the merge. In SPSS, value labels are attached automatically during the extraction process. Words in italics are meant to be replaced by the file or variable names that the user specifies.
Merging Restricted-use Data Files Using the School Control Number (CNTLNUMS)
When merging any of the school, principal, teacher, or school library media center files together for a given school, the school's control number, CNTLNUMS, is used to merge data files. The SPSS syntax is provided below.
get file = 'dataset1.sav';
sort cases by CNTLNUMS(A);
save outfile = 'dataset1.sav';
get file = 'dataset2.sav';
sort cases by CNTLNUMS(A);
save outfile = 'dataset2.sav';
match files file = 'dataset1.sav' * merges the two files and specifies dataset1 as
unit of analysis* /table 'dataset2' /by CNTLNUMS;
save outfile = 'mergeddatafile.sav'; *creates new merged filename*

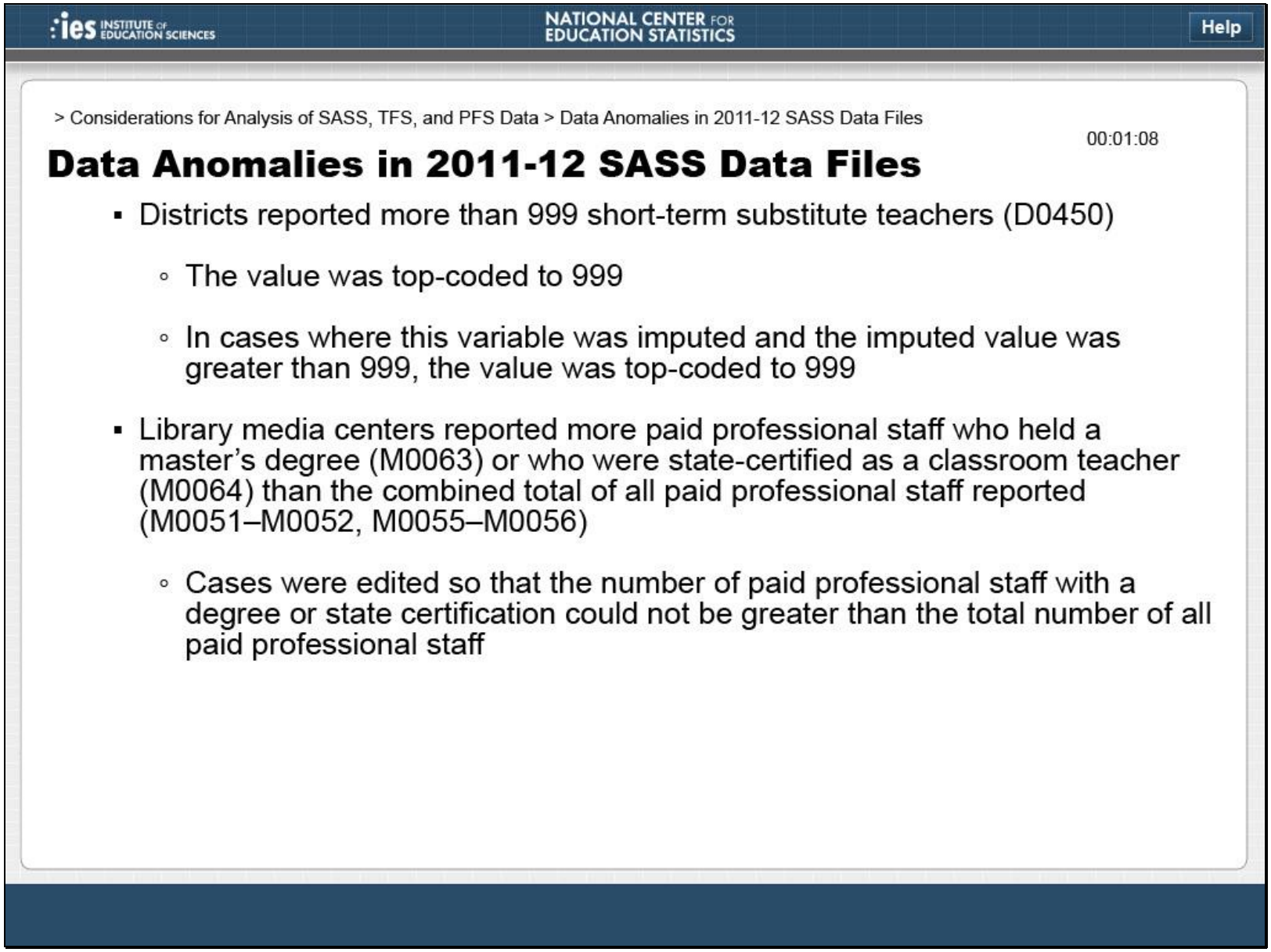

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Sample SPSS Syntax (Continued)
00:00:28
Sample SPSS Syntax for Merging Files Within SASS (Coninued)
get file = 'dataset1.sav';
sort cases by CNTLNUMS(A);
save outfile = 'dataset1.sav';
get file = 'dataset2.sav';
sort cases by CNTLNUMS(A);
save outfile = 'dataset2.sav';
match files file = 'dataset1.sav' * merges the two files and specifies dataset1 as unit of analysis* /table 'dataset2' /by CNTLNUMS;
save outfile = 'mergeddatafile.sav'; *creates new merged filename*
Merging Restricted-use Public School District Data File with Other SASS Public Sector Data Files
To merge the Public School District Data File with other public sector data files, the district's control number (CNTLNUMD) should be used.

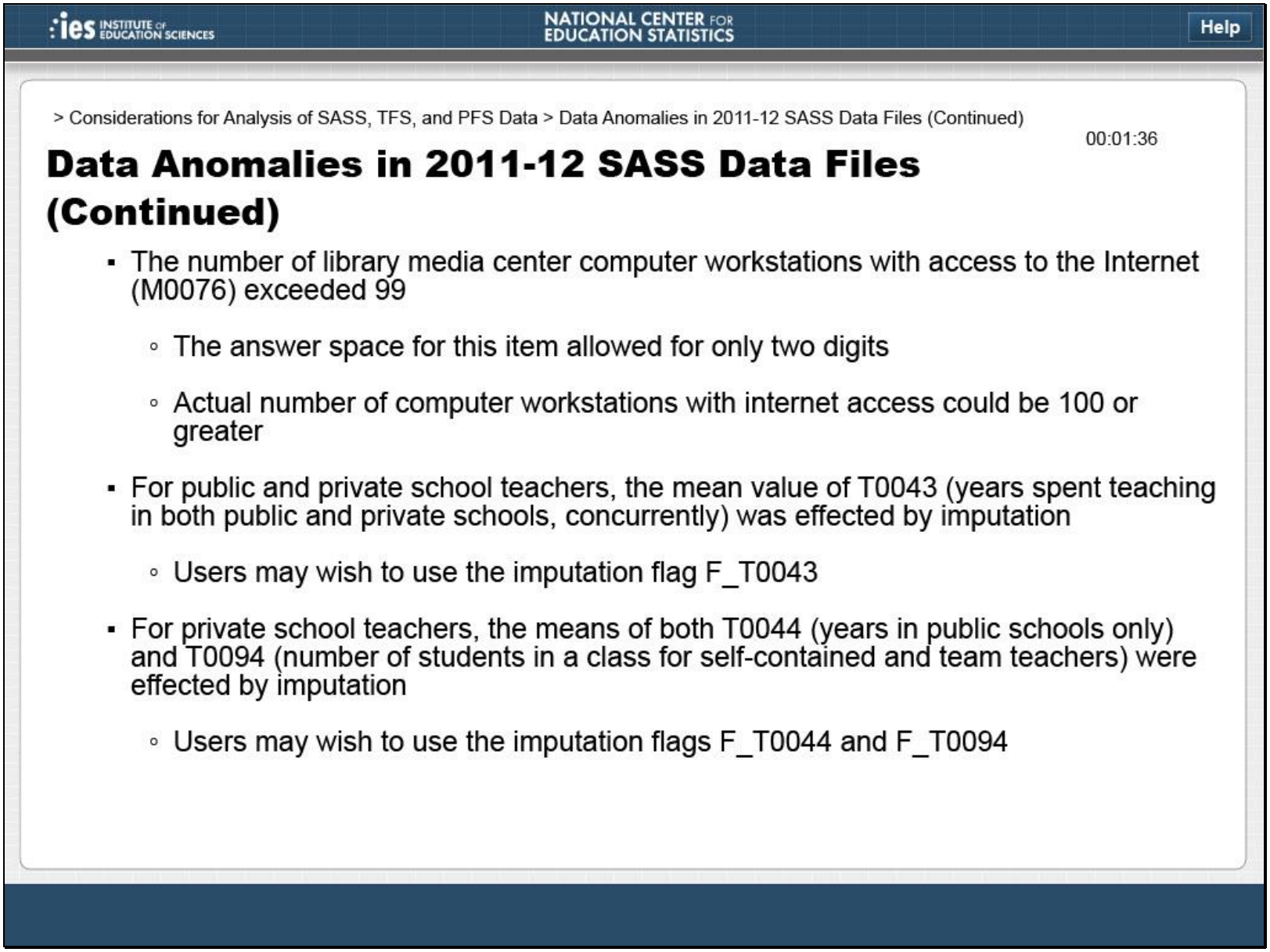

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Sample SAS Syntax
00:01:12
Sample SAS Syntax for Merging Data Files and Attaching Variable Labels
Merging Restricted-use Data Files Using the School Control Number (CNTLNUMS)
▪ Please note that both data files being merged must be sorted by the variable listed in the "by" statement prior to performing the merge.
▪ Comments to explain lines of code are contained within "/* */".
▪ Words in italics are meant to be replaced by the file or variable names that the user specifies.
proc sort data = dataset1;
by CNTLNUMS;
run;
proc sort data = dataset2;
by CNTLNUMS;
run;
data newfilename; /*creates new merged file name*/
merge dataset1 (in=a) dataset2; /* merges the two files and specifies dataset1 as
unit of analysis*/
by CNTLNUMS;
if a = 1; /*keeps all dataset1 records and only matching
dataset2 records*/
run;

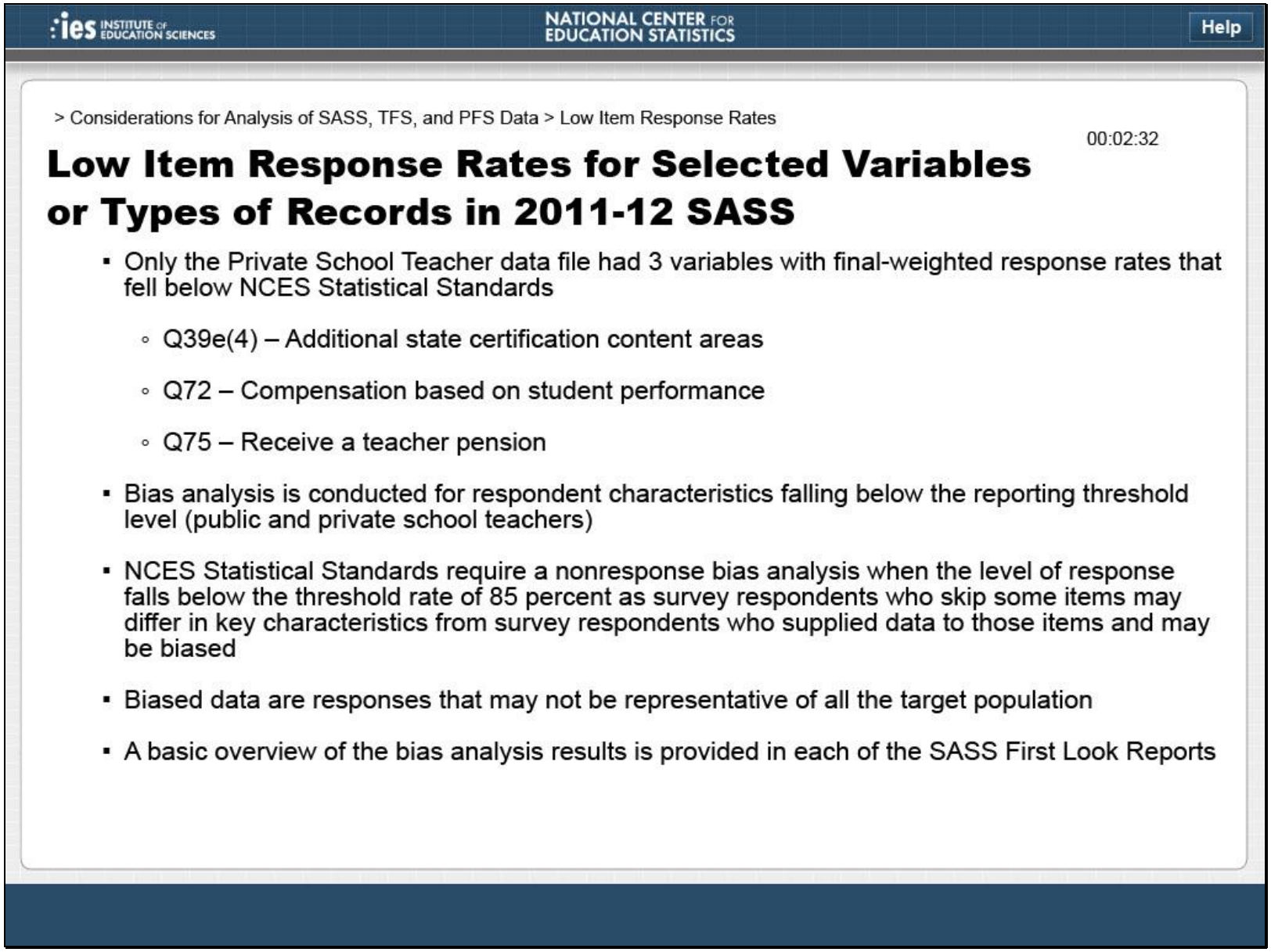

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Sample SAS Syntax (Continued)
00:01:09
Sample SAS Syntax for Merging Data Files and Attaching Variable Labels (Continued)
Merging the Restricted-use Public School District Data File with Other SASS Public Sector Data Files
To merge the Public School District Data File with other public sector data files, the district's control number (CNTLNUMD) should be used.
Value labels are not automatic in SAS
proc sort data = dataset1;
by CNTLNUMS;
run;
proc sort data = dataset2;
by CNTLNUMS;
run;
data newfilename; /*creates new merged file name*/
merge dataset1 (in=a) dataset2; /* merges the two files and specifies dataset1 as
unit of analysis*/
by CNTLNUMS;
if a = 1; /*keeps all dataset1 records and only matching
dataset2 records*/
run;

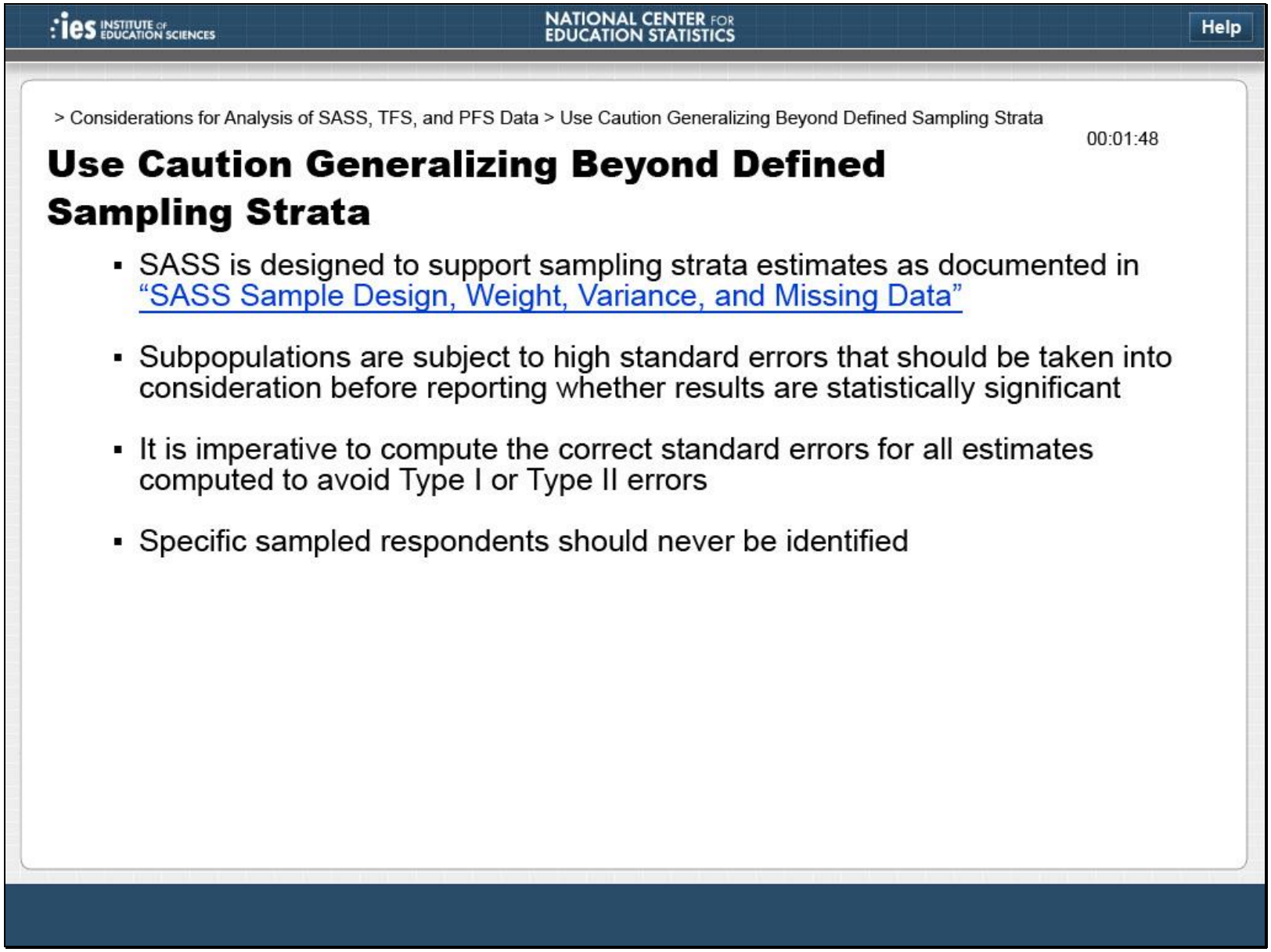

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Unit of Analysis
00:02:03
Unit of Analysis
Each data file in SASS has a different final weight
If you are only using one data file use the final weight in that file
Example: Public Teacher data file
Choose TFINWGT as your weight
If you have merged files, your merged dataset will have more than one final weight
Example: A merged school and teacher data file
Which weight should you choose?
It is extremely important to identify the appropriate unit of analysis and use the correct weight, or else your analytic approach may be misguided and have misleading results
Usually, considering the target population is the best way to select which weight variable to use
Considering which group the analysis is being generalized to represent will help you identify the unit of analysis for your research

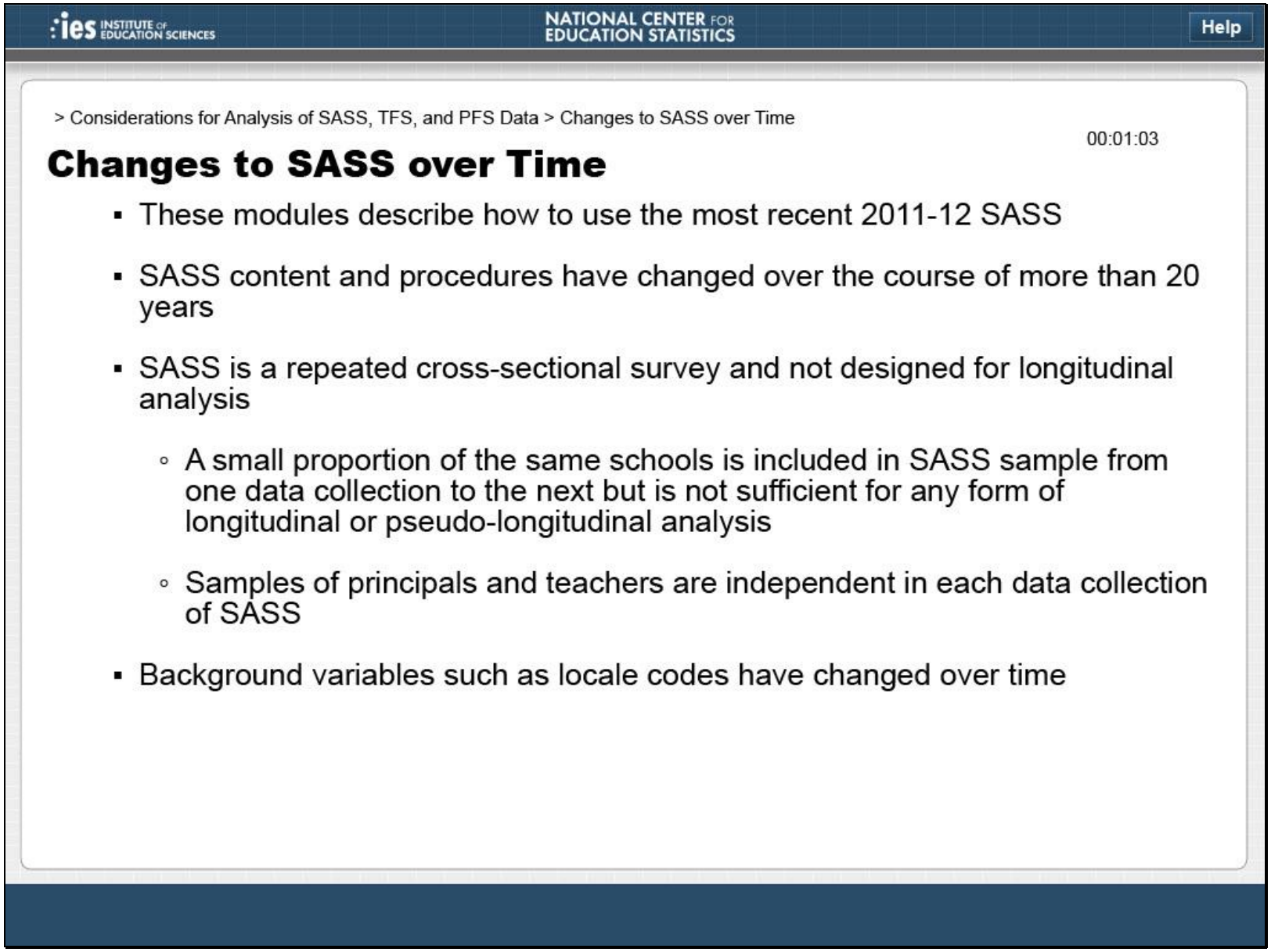

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Low Unit Response Rates in 2011-12 SASS Data Files
00:03:25
Low Unit Response Rates in 2011-12 SASS Data Files
NCES does not publish estimates when unit response rate is below 50 percent for a particular subgroup
Public School Principal: District of Columbia and Maryland had response rates below 50 percent
Private School Principal: Jewish principals had response rates below 50 percent
Public School: District of Columbia and Maryland had response rates below 50 percent
Private School: Jewish schools had response rates below 50 percent
School District: The single district in Hawaii did not respond to the school district survey
Library Media Center: District of Columbia and Maryland had response rates below 50 percent
Public School Teacher: City schools, Alaska, District of Columbia, Florida, Hawaii, Maryland and Rhode Island had response rates below 50 percent.
Private School Teacher: Teachers in other religious schools, nonsectarian schools, suburban schools, rural schools, combined schools, schools with less than 200 students, schools with more than 750 students, Baptist schools, Jewish schools, nonsectarian regular schools and nonsectarian special emphasis schools were below 50 percent

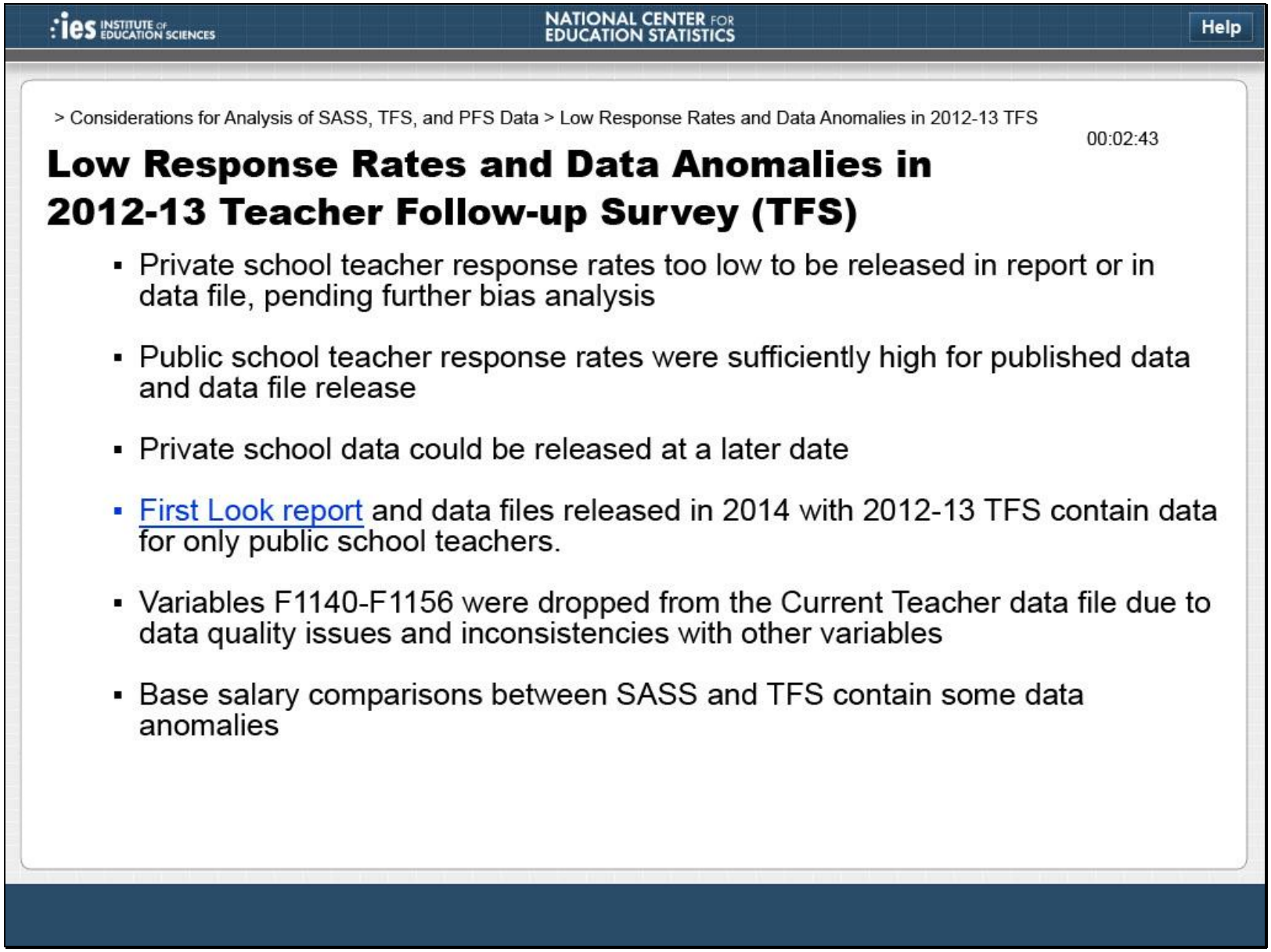

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Data Anomalies in 2011-12 SASS Data Files
00:01:08
Data Anomalies in 2011-12 SASS Data Files
▪ Districts reported more than 999 short-term substitute teachers (D0450)
◦ The value was top-coded to 999
◦ In cases where this variable was imputed and the imputed value was greater than 999, the value was top-coded to 999
▪ Library media centers reported more paid professional staff who held a master's degree (M0063) or who were state-certified as a classroom teacher (M0064) than the combined total of all paid professional staff reported (M0051–M0052, M0055–M0056)
◦ Cases were edited so that the number of paid professional staff with a degree or state certification could not be greater than the total number of all paid professional staff


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Data Anomalies in 2011-12 SASS Data Files (Continued)
00:01:36
Data Anomalies in 2011-12 SASS Data Files (Continued)
The number of library media center computer workstations with access to the Internet (M0076) exceeded 99
The answer space for this item allowed for only two digits
Actual number of computer workstations with internet access could be 100 or greater
For public and private school teachers, the mean value of T0043 (years spent teaching in both public and private schools, concurrently) was effected by imputation
Users may wish to use the imputation flag F_T0043
For private school teachers, the means of both T0044 (years in public schools only) and T0094 (number of students in a class for self-contained and team teachers) were effected by imputation
Users may wish to use the imputation flags F_T0044 and F_T0094


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Low Item Response Rates
00:02:32
Low Item Response Rates for Selected Variables or Types of Records in 2011-12 SASS
Only the Private School Teacher data file had 3 variables with final-weighted response rates that fell below NCES Statistical Standards
Q39e(4) – Additional state certification content areas
Q72 – Compensation based on student performance
Q75 – Receive a teacher pension
Bias analysis is conducted for respondent characteristics falling below the reporting threshold level (public and private school teachers)
NCES Statistical Standards require a nonresponse bias analysis when the level of response falls below the threshold rate of 85 percent as survey respondents who skip some items may differ in key characteristics from survey respondents who supplied data to those items and may be biased
Biased data are responses that may not be representative of all the target population
A basic overview of the bias analysis results is provided in each of the SASS First Look Reports


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Use Caution Generalizing Beyond Defined Sampling Strata
00:01:48
Use Caution Generalizing Beyond Defined Sampling Strata
SASS is designed to support sampling strata estimates as documented in "SASS Sample Design, Weight, Variance, and Missing Data"
Subpopulations are subject to high standard errors that should be taken into consideration before reporting whether results are statistically significant
It is imperative to compute the correct standard errors for all estimates computed to avoid Type I or Type II errors
Specific sampled respondents should never be identified

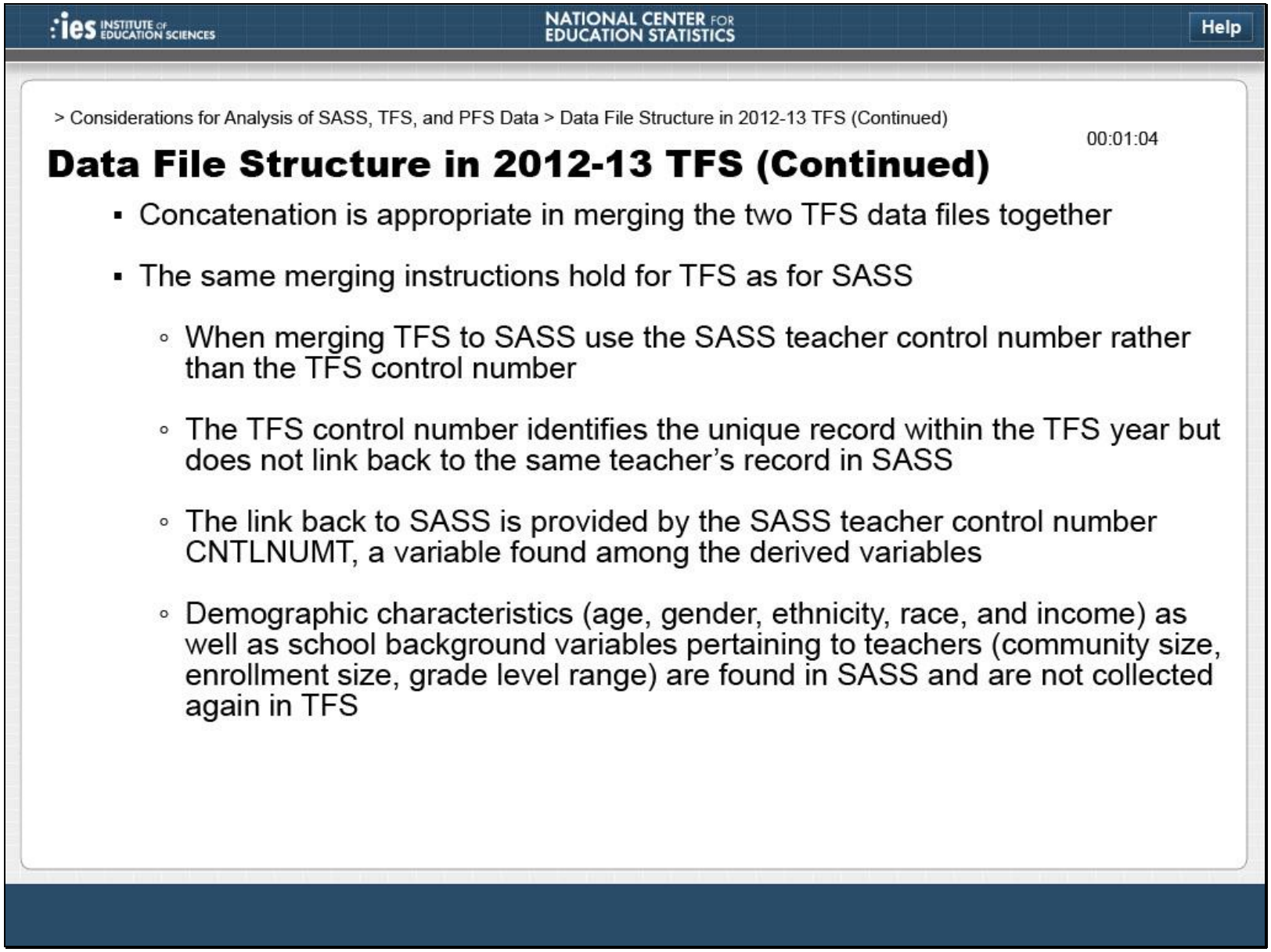

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Changes to SASS over Time
00:01:03
Changes to SASS over Time
These modules describe how to use the most recent 2011-12 SASS
SASS content and procedures have changed over the course of more than 20 years
SASS is a repeated cross-sectional survey and not designed for longitudinal analysis
A small proportion of the same schools is included in SASS sample from one data collection to the next but is not sufficient for any form of longitudinal or pseudo-longitudinal analysis
Samples of principals and teachers are independent in each data collection of SASS
Background variables such as locale codes have changed over time

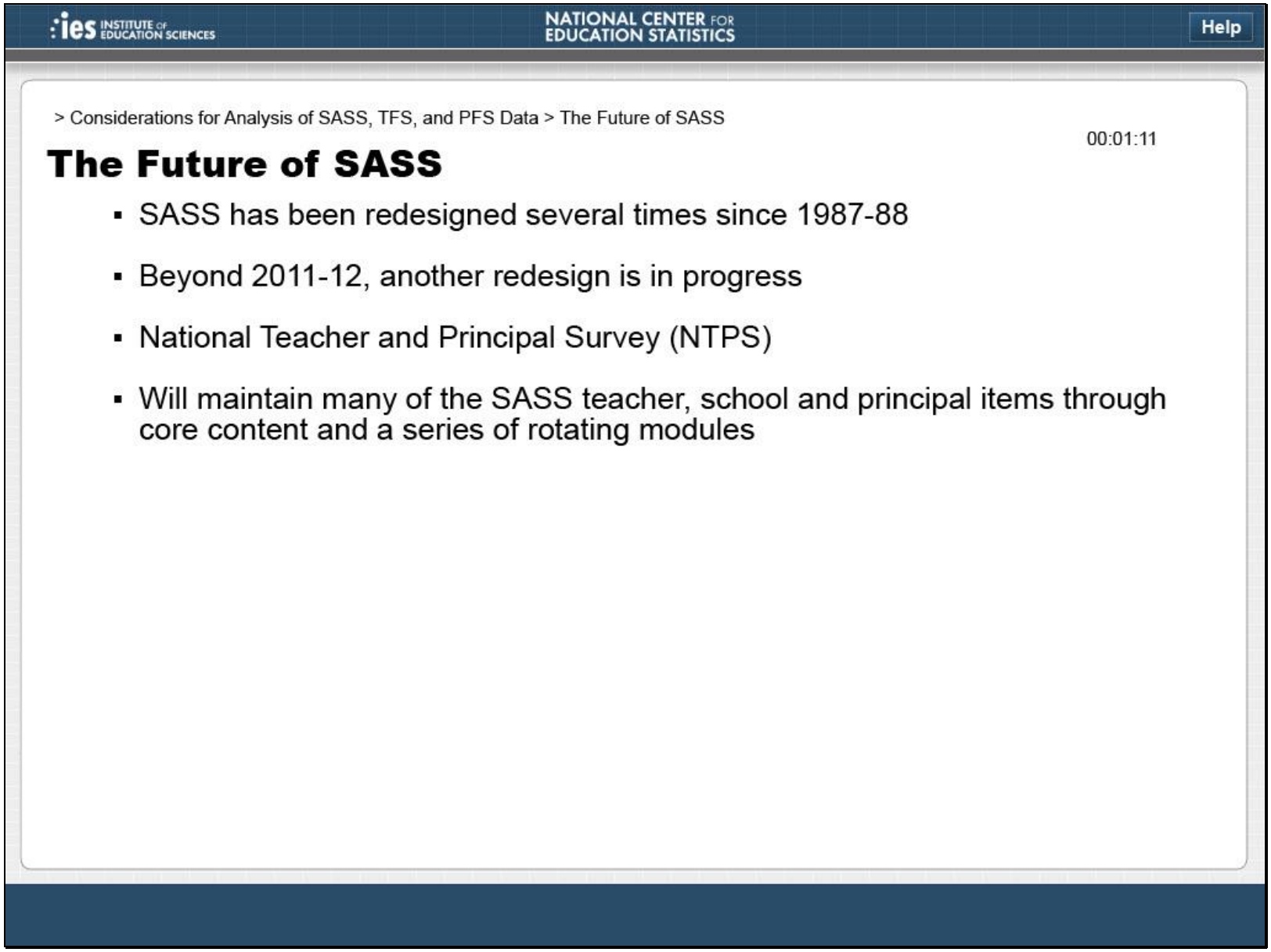

ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Low Response Rates and Data Anomalies in 2012-13 TFS
00:02:43
Low Response Rates and Data Anomalies in 2012-13 Teacher Follow-up Survey (TFS)
Private school teacher response rates too low to be released in report or in data file, pending further bias analysis
Public school teacher response rates were sufficiently high for published data and data file release
Private school data could be released at a later date
First Look report and data files released in 2014 with 2012-13 TFS contain data for only public school teachers.
Variables F1140-F1156 were dropped from the Current Teacher data file due to data quality issues and inconsistencies with other variables
Base salary comparisons between SASS and TFS contain some data anomalies

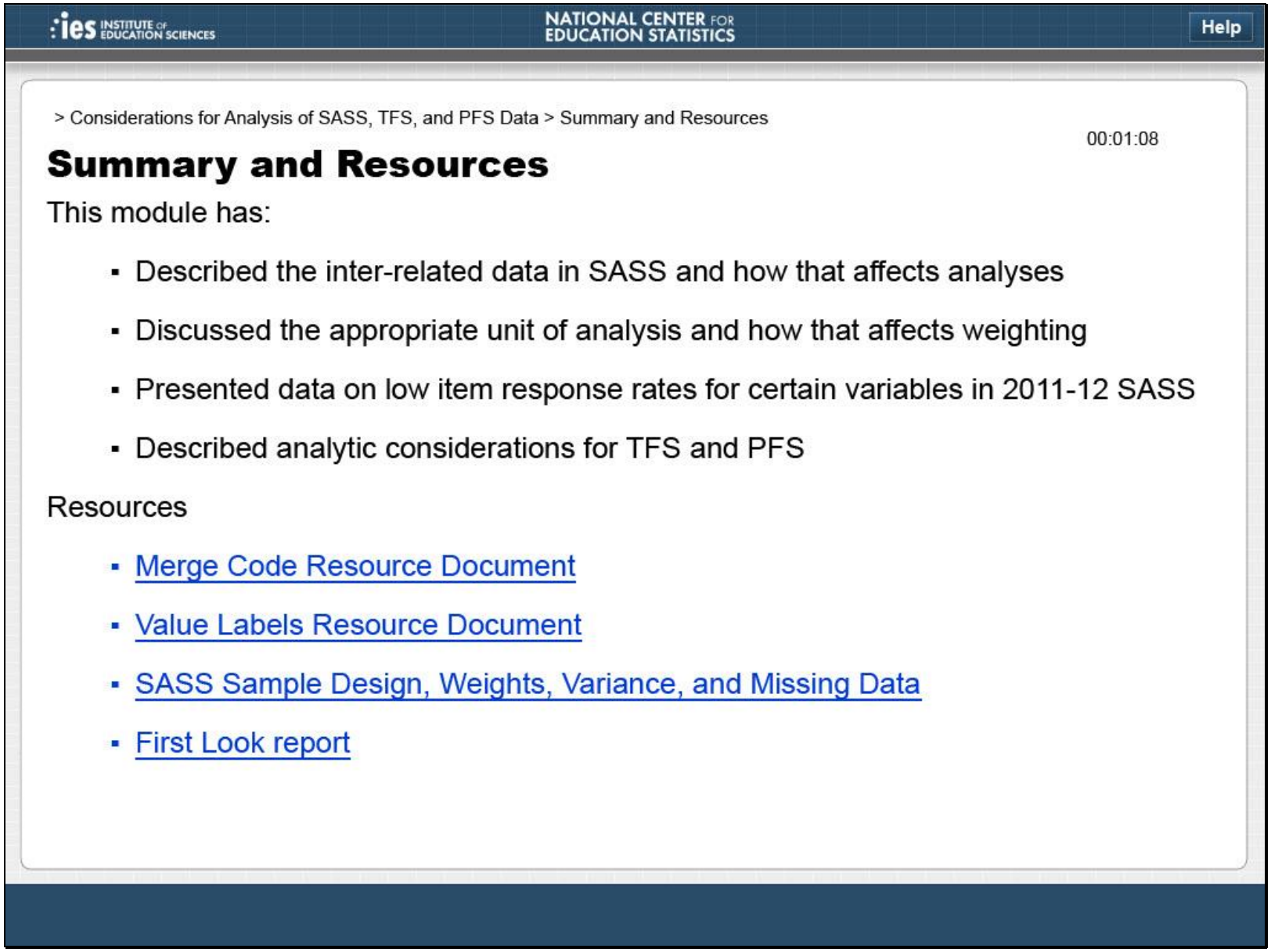
00:01:07

Data File Structure in 2012-13 TFS

- Two files: Current Teacher and Former Teacher
- TFS data files contain SASS variables for the TFS sample respondents
- Similar to SASS, some analyses may require merging files


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Data File Structure in 2012-13 TFS (Continued)
00:01:04
Data File Structure in 2012-13 TFS (Continued)
▪ Concatenation is appropriate in merging the two TFS data files together
▪ The same merging instructions hold for TFS as for SASS
◦ When merging TFS to SASS use the SASS teacher control number rather than the TFS control number
◦ The TFS control number identifies the unique record within the TFS year but does not link back to the same teacher's record in SASS
◦ The link back to SASS is provided by the SASS teacher control number CNTLNUMT, a variable found among the derived variables
◦ Demographic characteristics (age, gender, ethnicity, race, and income) as well as school background variables pertaining to teachers (community size, enrollment size, grade level range) are found in SASS and are not collected again in TFS


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > The Future of SASS
00:01:11
The Future of SASS
SASS has been redesigned several times since 1987-88
Beyond 2011-12, another redesign is in progress
National Teacher and Principal Survey (NTPS)
Will maintain many of the SASS teacher, school and principal items through core content and a series of rotating modules


ies INSTITUTE OF EDUCATION SCIENCES
NATIONAL CENTER FOR EDUCATION STATISTICS
Help
> Considerations for Analysis of SASS, TFS, and PFS Data > Summary and Resources
00:01:08
Summary and Resources
This module has:
Described the inter-related data in SASS and how that affects analyses
Discussed the appropriate unit of analysis and how that affects weighting
Presented data on low item response rates for certain variables in 2011-12 SASS
Described analytic considerations for TFS and PFS
Resources
Merge Code Resource Document
Value Labels Resource Document
SASS Sample Design, Weights, Variance, and Missing Data
First Look report